AD_____

Award Number: W81XWH-05-1-0138

TITLE: Identifying Novel Drug Targets for the Treatment of Tuberous Sclerosis Complex Using High Throughput Technologies

PRINCIPAL INVESTIGATOR: David Sabatini, M.D., Ph.D.

CONTRACTING ORGANIZATION: Whitehead Institute for Biomedical Research
Cambridge, MA 02142-1479

REPORT DATE: July 2007

TYPE OF REPORT: Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-07-2007	Final Addendum	8 Dec 2005 – 7 Jun 2007

4. TITLE AND SUBTITLE

Identifying Novel Drug Targets for the Treatment of Tuberous Sclerosis Complex Using High Throughput Technologies

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0138

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

David Sabatini, M.D., Ph.D.

E-Mail: anne@broad.mit.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Whitehead Institute for Biomedical Research
Cambridge, MA 02142-1479

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES – Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

In a patient with Tuberous Sclerosis Complex (TSC), the problematic cells that initiate and constitute tumors have lost TSC1 or TSC2 function. A promising approach for treatment would be to target members of the pathway with which TSC1/2 proteins interact. In cultured drosophila cells, we proposed to rapidly identify genes whose RNAi-mediated reduction in expression (1) Prevents growth/proliferation of TSC1 or TSC2-deficient cells without affecting normal cells. (2) Induces apoptosis/cell death in TSC1 or TSC2-deficient cells without killing normal cells. (3) Reverts TSC1 or TSC2-deficient cells to a normal phenotype, as determined by measuring a reporter of cell growth pathway activation and cell morphology. We have (1) advanced genome-wide RNA interference living cell microarrays from proof-of-principle to a robust technology, (2) developed software to analyze these screens, a previously formidable challenge, and (3) completed genome-wide experiments to identify genes involved in the TSC pathway.

15. SUBJECT TERMS
Genome-wide, RNA interference, TSC

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	22	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

In a patient with Tuberous Sclerosis Complex (TSC), the problematic cells that initiate and constitute tumors have lost TSC1 or TSC2 function. A promising approach to develop an effective small-molecule drug for TSC would be to target members of the pathway with which TSC1/2 proteins interact. Such a drug could either bypass the requirement for TSC1/2 or specifically kill or arrest cells that have lost TSC function. Unfortunately, very little is known about the proteins in this pathway or how they interact, so we cannot make an educated guess about which proteins should be targeted to kill, arrest, or revert TSC mutant cells. We are aware of significant gaps in our knowledge of the proteins in this pathway and how they interact[1]. Even for those proteins which are already known to be in the pathway[2], predicting the effects on living cells of targeting those proteins is unreliable. Clearly, we need to better understand this pathway in order to more rationally design drugs to treat TSC. These patients could likely be successfully treated with small molecule drugs if drug targets could be identified which cause only cells deficient in TSC1 or TSC2 function to arrest, die, or revert to normal without disrupting the patient's normal remaining cells. In cultured drosophila cells, we proposed to rapidly identify genes whose RNAi-mediated reduction in expression (1) Prevents growth/proliferation of TSC1 or TSC2-deficient cells without affecting normal cells. (2) Induces apoptosis/cell death in TSC1 or TSC2-deficient cells without killing normal cells. (3) Reverts TSC1 or TSC2-deficient cells to a normal phenotype, as determined by measuring a reporter of cell growth pathway activation and cell morphology.

Body

After the no-cost extension granted, we have now completed the goals outlined in the original research proposal:

Task 1: Prepare for genome-wide screens:

Print the genome-wide RNAi microarrays using our existing libraries of dsRNA

We have developed the dsRNA printing technology and scaled it up to a higher-throughput, more robust format to allow genome-wide screens. Previously, we had printed a maximum of 384 genes on a single slide[3]. We have now successfully printed ~20,000 dsRNAs, covering the vast majority of Drosophila genes, onto four glass microscope slides. This printing has been developed at a feature density which allows 5600 genes to be spotted per slide.

We overcame technical difficulties relating to the unwanted spreading of dsRNA in the spots and the imprecision of the DNA microarraying robot. We also were able to train new personnel to successfully carry out the experiments from printing, to cell seeding to image acquisition and analysis. The transferability of the techniques to new personnel indicates that this technology is robust enough to also be transferred to other laboratories. Since the extension was granted, we have printed dozens of sets of these genome-wide slides and have successfully used them for several screens.

Optimize software for these cells and these assays and write code to improve data extraction

Image cytometry (automated cell image analysis) simultaneously measures many valuable features of cells: the intensity, texture and localization of each fluorescently labeled cellular component (e.g. DNA or protein) within each subcellular compartment, as well as the number, size, and shape of each subcellular compartment. This type of full phenotypic analysis is necessary for our aim of identifying reversion to a normal phenotype. Drosophila cells were notoriously difficult to identify in images[4] using existing software. In addition, our project required the accurate measurement of a large number of cellular features, many of which were not measurable using commercial software.

Our laboratory therefore initiated an open-source software project, CellProfiler, to address these substantial challenges (Figure 1 and 2). CellProfiler allows accurate quantitative measurement of many cellular properties, including cell count, cell size, cell cycle distribution, organelle size, and the levels and localization of proteins and phospho-proteins. The software is user-friendly, flexible, modular, open-source, and free, making it a useful tool to share, compare, test, adapt, and further develop image analysis methods in the scientific community.

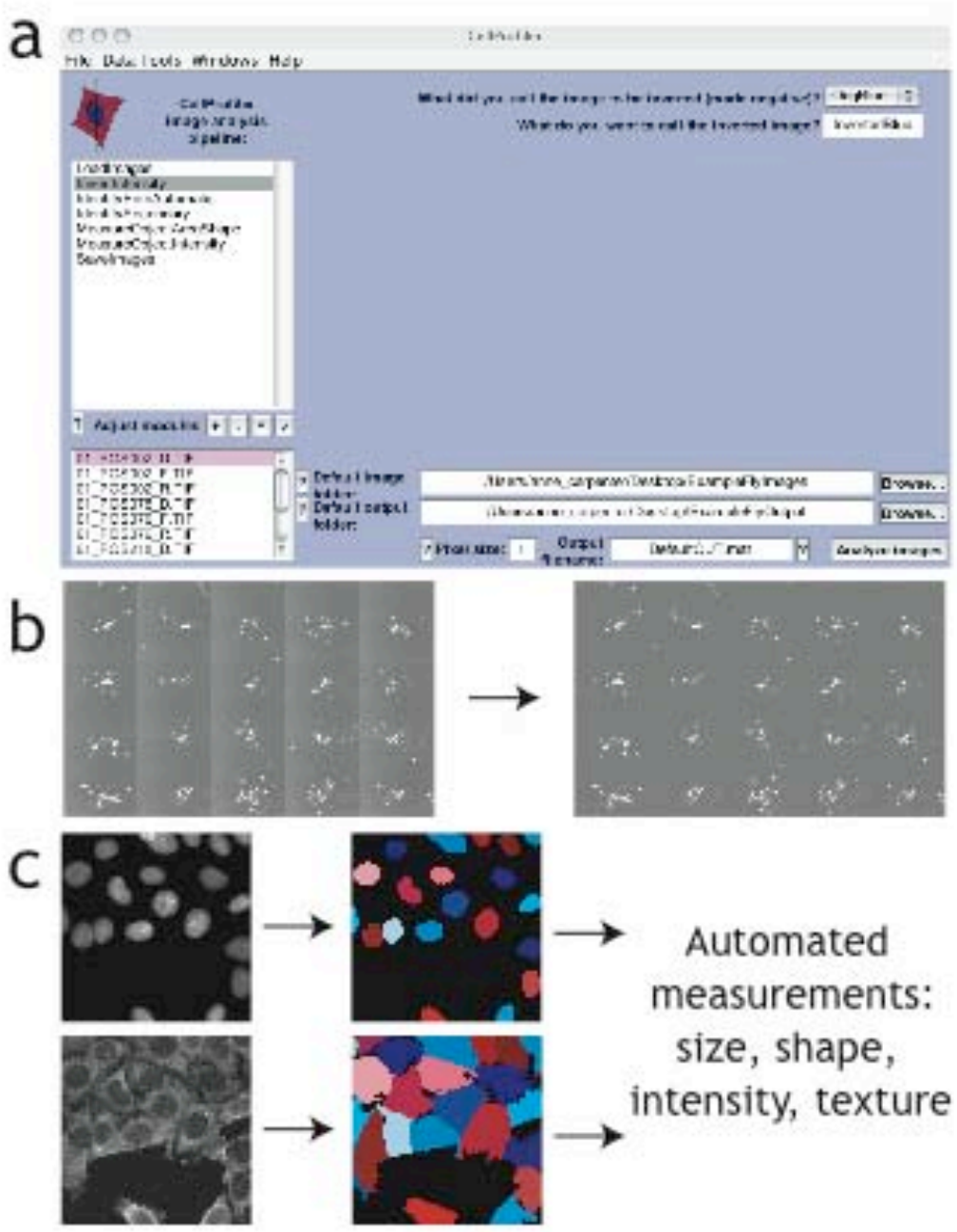

Figure 1: a. Main CellProfiler interface, with a simple analysis pipeline displayed. b. Image processing example: Uneven illumination of the field of view is noticeable when nearby images are placed adjacent to each other and can significantly affect intensity measurements from uncorrected images (left). The left side of each image in this 3 row, 5 column tiled image is brighter than the right side. CellProfiler's CorrectIllumination_Calculate and CorrectIllumination_Apply modules correct these anomalies (right). Images were brightness- and contrast-enhanced to display this effect. c. In typical usage, CellProfiler identifies nuclei first and cell edges are identified surrounding each nucleus using a cell stain image. Measurements are then made.

During the original project period, CellProfiler was adapted and optimized for the research projects proposed in several ways:

1. Algorithms for the successful identification of Drosophila nuclei and cells in images were developed and validated[5].
2. We added the ability to measure a large number of sophisticated measurements to CellProfiler, including many measures of the size, shape, intensity, and texture of cells.
3. CellProfiler was validated for many key phenotypes (examples, Figure 2). We tested its ability to measure cell number, cell size, cell cycle distribution (based on DNA content), and amount of fluorescence per cell using cells with known variations in these features.
4. The user interface of CellProfiler was improved dramatically to eliminate some of the tedium of high-throughput image analysis and to allow non-experts the ability to conduct image analysis experiments, including documentation and a manual.
5. CellProfiler was adapted to make use of a cluster of computers, so that it can analyze images at a pace faster than image acquisition.
6. CellProfiler was adapted to export measurements to Excel and also to a database, a necessary feature for large-scale genome-wide experiments.
7. CellProfiler was beta-tested by several academic and industry research groups.
8. CellProfiler was released for free to the public, allowing further development by the open-source community.

During the extension period, we wrote up and published validation of CellProfiler in a high-profile publication[6].

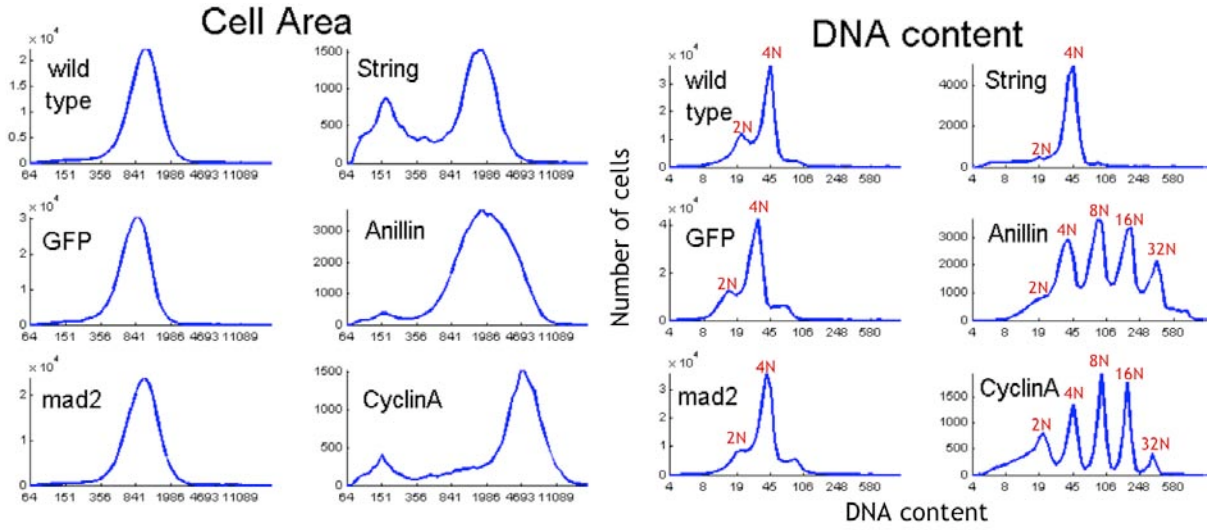

Figure 2: Images from slides with the indicated gene knocked down by RNA interference were analyzed using CellProfiler software to create histograms of cell area (left) and DNA content (right). The results are consistent with existing literature, validating the software.

In addition, we initiated another software project, CellVisualizer, to allow visualization and extraction of the measurements made by CellProfiler (Figure 3). Because the number and richness of measurements from image-based genome-wide assays are unprecedented, other fields must advance to accommodate this influx of data. A full

phenotypic analysis of 20,000 images, each containing about 400 cells, includes ~8 million cells and produces roughly three billion measurements (400 measures per cell, including size, shape, and the intensity and texture of three fluorescent stains). CellVisualizer allows biologists without database experience to analyze data from genome-wide screens, including identifying unusual samples based on their quantitative measurements and viewing the original images corresponding to those samples.

We have discovered that these developments in image analysis and data visualization have been critical for our ability to accurately identify genes of interest in the large-scale assays relating to the aims of this proposal.

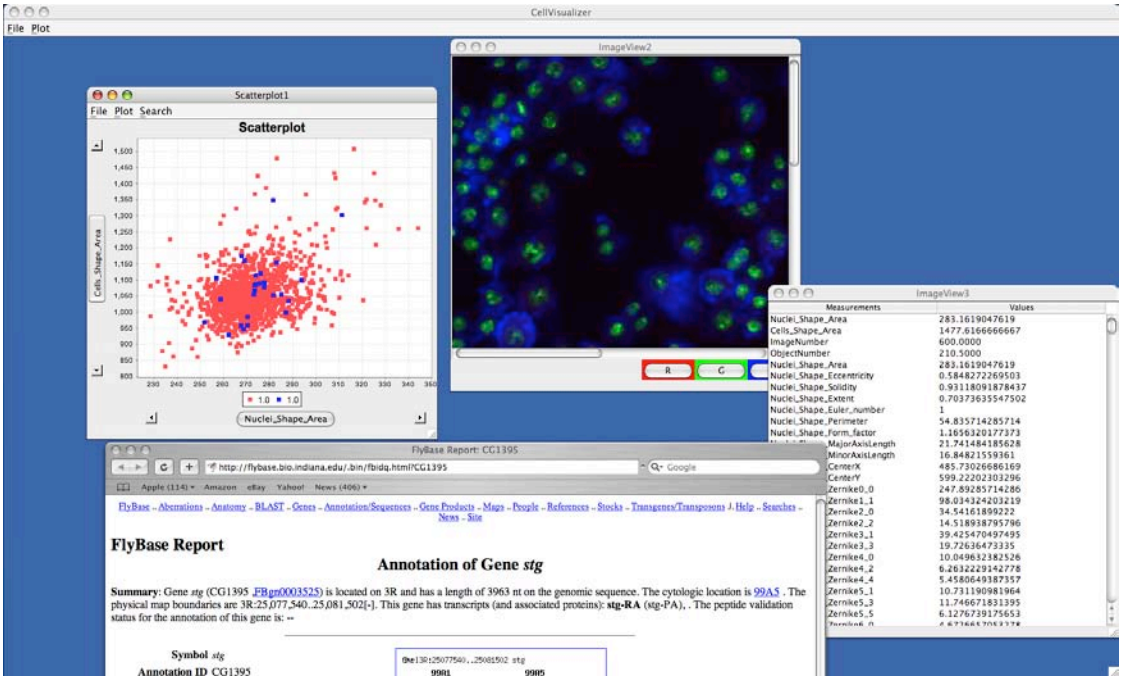

Figure 3: We developed CellVisualizer, shown here, to allow visualization of data from genome-wide screens. Genes are shown in a scatterplot based on measurements of cells from images with that gene knocked down (top left). Clicking a gene in the scatterplot allows viewing the raw, original image (top middle), the numerical measurements from that image (right) and links to the external gene database FlyBase (bottom).

During the extension period, importantly, we have developed machine-learning methods and incorporated them into CellVisualizer, allowing complex phenotypes to be scored using morphology markers (DNA and actin stain). In this approach, biologists train a computer to recognize cellular phenotypes of interest (often quite subtle) in fluorescence microscopy images (Figure 4). Machine learning algorithms then distinguish cells of interest based on each cell's cytological profile, its rich set of image cytometry-measured features including size, shape, intensity and texture. This approach allows rapid scoring of millions of individual cells.

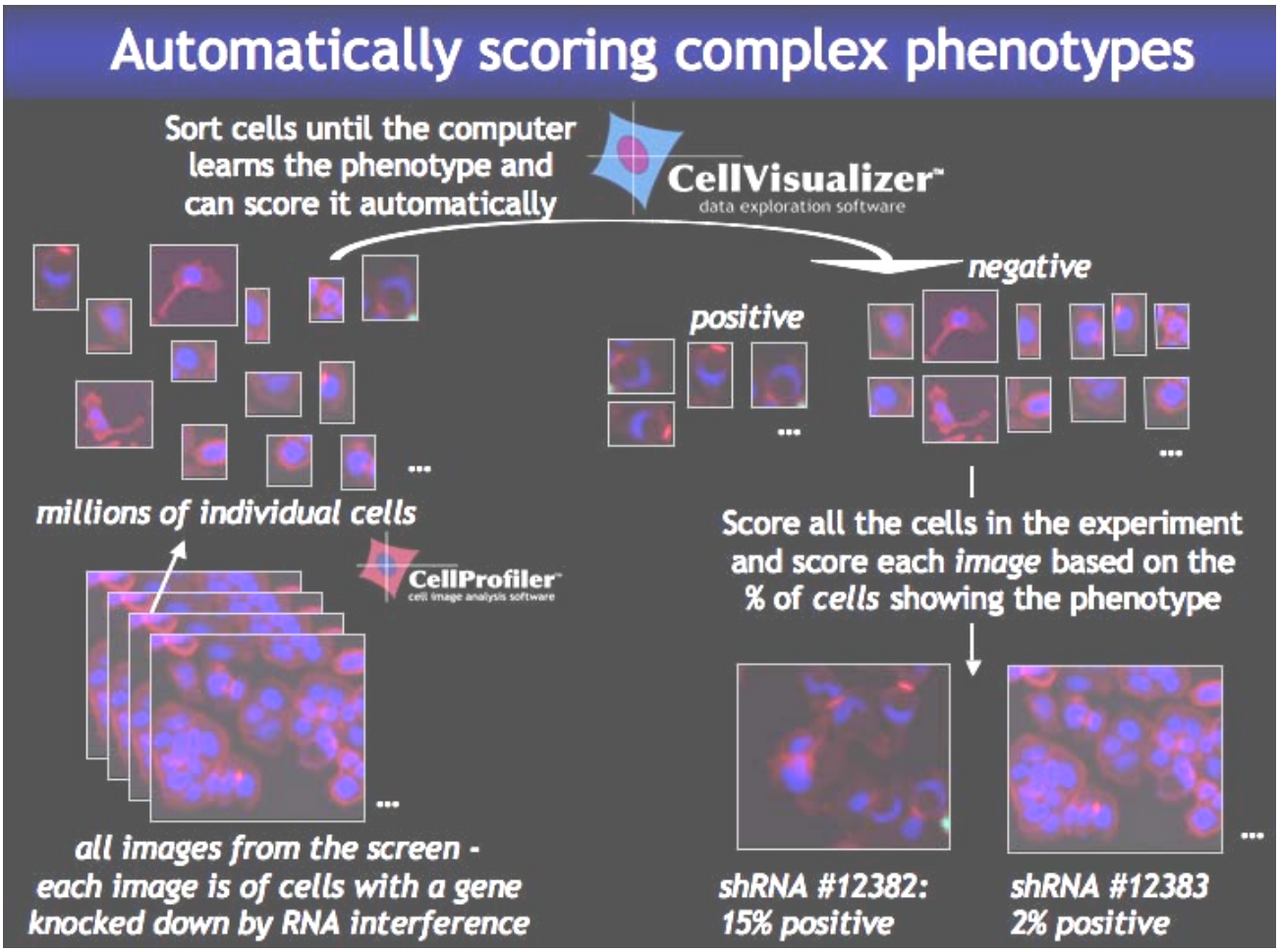

Figure 4: Machine learning methods in CellVisualizer allow rapid scoring of phenotypes from images in genome-wide screens. Shown here is scoring for a nuclear morphology-related phenotype, crescent-shaped nuclei.

Task 2: Conduct genome-wide screens:

Prepare RNAi/cell microarrays in duplicate (Total = 24 microscope slides)

During the original project period, we successfully printed, seeded, fixed, stained, imaged, and analyzed genome-wide cell microarrays. In particular, we conducted the following genome-wide experiments:

1. wild type drosophila Kc167 cells treated with dsRNA against nearly every gene in the Drosophila genome (quadruplicate).
2. drosophila S2 cells expressing a modified S6 reporter protein (a reporter of a branch of the cell growth pathway), stained for phospho-S6 (duplicate + rapamycin, duplicate - rapamycin). Note: RNA interference was not very effective on this set of slides.

As of August 2006, we took a step back to smaller-scale experiments in an attempt to work out these techniques more robustly. At that time, we had completed some of the original project goals, having completed the following:

1. Screens covering all the kinases and phosphatases (288 genes) in Kc167 cells that are:
    - Wild type
    - TSC1 knockdown
    - TSC2 knockdown
2. The first genome-wide experiments like (1) had been imaged for wild type cells.
3. The first kinase and phosphatase screens in S6 wild-type cells had been prepared and imaged.

Now finally, during the extension period, we completed the work described in the Statement of Work in the research proposal, albeit with modifications as described here:

Task 2a: Prepare the following RNAi/cell microarrays in duplicate (Total = 24 microscope slides)

1. Cells: wild type drosophila Kc cells bathed in dTSC2 RNA interference reagent
   Spots: each gene in the drosophila genome
   Staining: Hoechst 33342 (DNA), FITC-phalloidin (actin), and Sytox dye (dead cells)

We have completed these screens, in triplicate rather than duplicate, to improve data quality. We have omitted the Sytox dye because we decided that cell count would sufficiently indicate samples where cells had died (in addition to samples where cells failed to grow at a normal rate) and that these two cases could be easily separated in downstream followup work. Results are described below.

2. Same as (1) except cells will be bathed in dTSC1 RNA interference reagent

We have completed this screen. Results are described below.

3. Cells: drosophila Kc cells expressing human S6 protein, bathed in dTSC2 RNA interference reagent
   Spots: each gene in the drosophila genome
   Staining: Hoechst 33342 (DNA), FITC-phalloidin (actin), and anti-phospho-S6 antibody (a reporter of a branch of the cell growth pathway)

We completed this screen using wild type cells rather than TSC1 or TSC2 knockdown cells because it proved to be more efficient to screen wild type cells and then follow up the hits from that genome-wide screen, in more detail. We have therefore done more extensive follow up on each gene to make up for the lack of a genome-wide screen at the primary screening stage. Results are described below.

4. Same as (3) except cells will be bathed in dTSC1 RNA interference reagent

See description for Task 2a.3.

Task 2b: Collect images using automated microscope

Completed. Results are described below.

Task 2c: Analyze images for changes in cell growth, proliferation, and apoptosis/death and for changes in activation of S6 and reversion to normal morphology.

Completed, except S6 scaled to a secondary screen as described in 2a.3. Results are described below.

Results from Task 2:

We have performed the first genome-wide RNAi screens on the TSC1/2 and TOR pathways. In each screen, we tested a library of 20,000 dsRNAs targeting over 13,000 *Drosophila* genes. The results are organized here based on the original three Aims of the proposal:

Aim 1: Identify genes whose corresponding RNAi's prevent growth/proliferation of TSC1 or TSC2-deficient cells without affecting normal cells.

Aim 2: Identify genes whose corresponding RNAi's induce apoptosis/cell death in TSC1 or TSC2-deficient cells without killing normal cells.

We identified genes that reduce the cell count in TSC1 or TSC2-deficient cells without substantially affecting the proliferation of normal cells (Figure 5). This will include genes that induce apoptosis/cell death in TSC1 or TSC2-deficient cells without substantially affecting the apoptotic rate of normal cells.

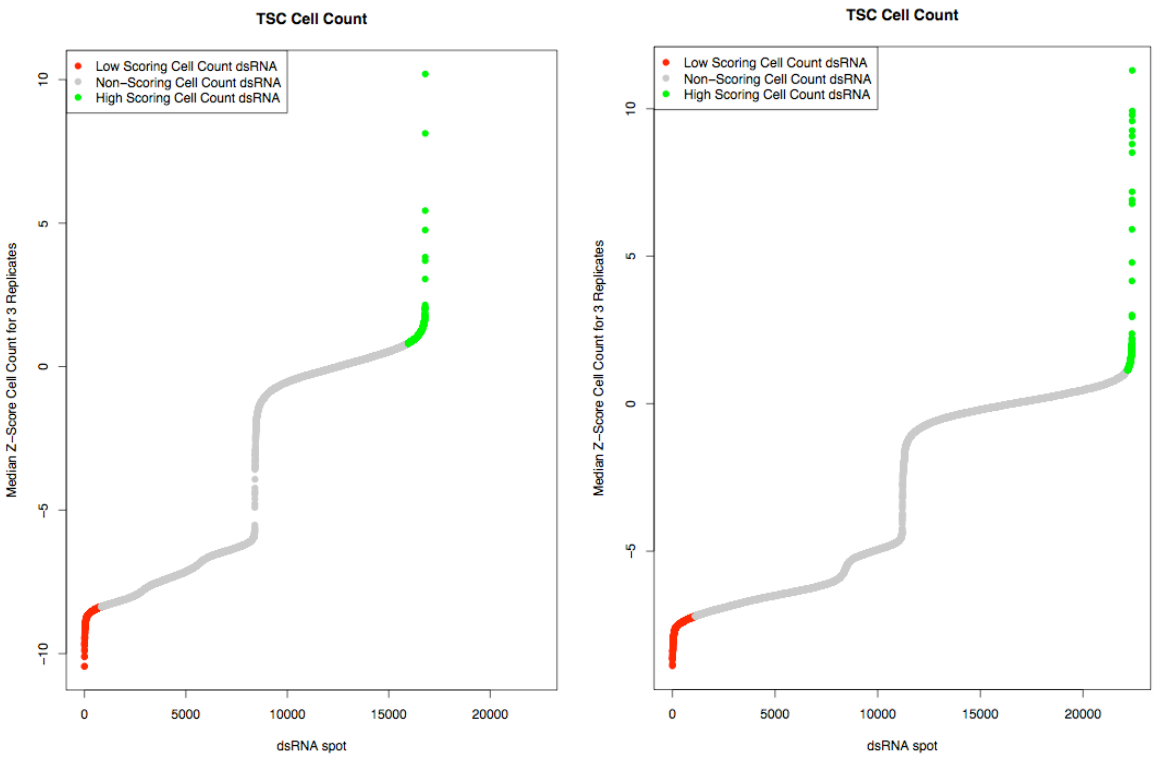

Figure 5: Overview of genome-wide results for cell count (proliferation) screen. TSC1 (left) and TSC2 (right). Low hits are of interest. These hits are then further screened, requiring that they show "normal" cell count in wild-type conditions, defined as between the 10th and 90th percentile of cell counts for the control samples (GFP samples). The top 6 genes are shown in the Table of hits below.

TSC1 Cell count hits:

DIAP1 Inhibitor of apoptosis
CG1883 Ribosomal protein S7 is referred to in FlyBase by the symbol RpS7 (CG1883, FBgn0039757)
CG5271 Ribosomal protein S27A is referred to in FlyBase by the symbol RpS27A (CG5271, FBgn0003942)
CG4897 Ribosomal protein L7 is referred to in FlyBase by the symbol RpL7 (CG4897, FBgn0005593)
CG5844
CG4111 Ribosomal protein L35 is referred to in FlyBase by the symbol RpL35 (CG4111, FBgn0029785)

TSC2 Cell Count hits:

CG2331 TER94 (CG2331, FBgn0024923)
CG14792 stubarista, sta (CG14792, FBgn0003517)
CG8615 Ribosomal protein L18 is referred to in FlyBase by the symbol RpL18 (CG8615, FBgn0035753)
CG4918 Ribosomal protein LP2 is referred to in FlyBase by the symbol RpLP2 (CG4918, FBgn0003274)
CG2960 Ribosomal protein L40 is referred to in FlyBase by the symbol RpL40 (CG2960, FBgn0003941)
DIAP1 Inhibitor of apoptosis

Clearly, it will be of interest to determine more carefully whether ribosomal genes' knockdown produces a general cell proliferation defect or whether it is truly TSC1/2 specific. The unannotated genes are particularly interesting to pursue.

Aim 3: Identify genes whose corresponding RNAi's revert TSC1 or TSC2-deficient cells to a normal phenotype, as determined by measuring (a) cell morphology and (b) a reporter of cell growth pathway activation.

Towards Aim 3 (a), measuring reversion of cell morphology, we settled on measurement on the cell size phenotype as the most dramatic TSC1 and TSC2 knockdown induced change in morphology. TSC1 and TSC2 knockdown by RNAi both cause the cells to be larger than normal. We identified genes that revert this phenotype while causing relatively little change in wild type cells (Figure 6).

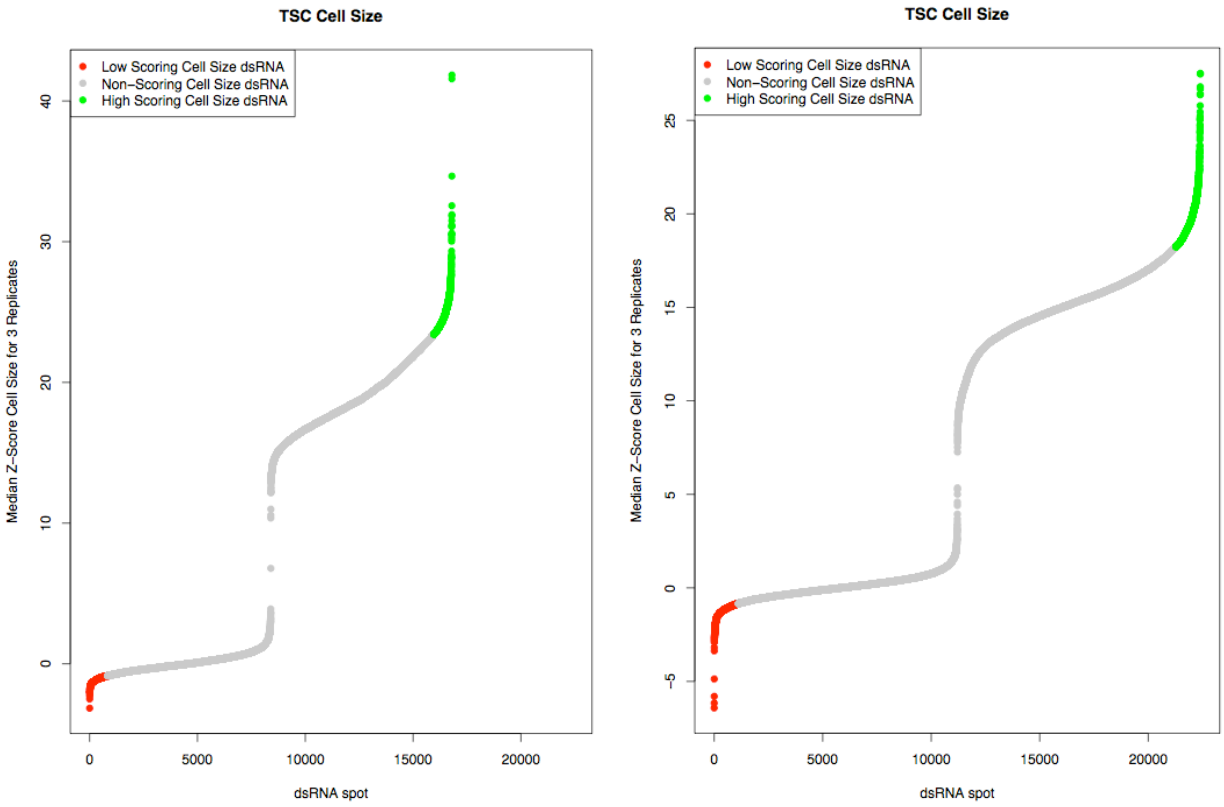

Figure 6: Overview of genome-wide results for cell size (morphology) screen. TSC1 (left) and TSC2 (right). Low hits are of interest. These hits are then further screened, requiring that they show "normal" cell size in wild-type conditions, defined as between the 10th and 90th percentile of cell size for the control samples (GFP samples). The top 6 genes are shown in the Table of hits below.

TSC1 Cell size hits:

CG1451 APC-like is referred to in FlyBase by the symbol Apc (CG1451, FBgn0015589)
CG17369 Vacuolar H+-ATPase 55kD B subunit is referred to by the symbol Vha55 (CG17369, FBgn0005671)
CG30035, CG7797
CG4688
CG12688
CG30396 Gustatory receptor 58a is referred to in FlyBase by the symbol Gr58a (CG30396, FBgn0041239)

TSC2 Cell size hits:

CG7457
CG2102 The gene castor is referred to in FlyBase by the symbol cas (CG2102, FBgn0004878)
CG5880
CG16905
CG8169 This gene is referred to in FlyBase by the symbol Pms2 (CG8169, FBgn0011660)
CG9369 The gene miniature is referred to in FlyBase by the symbol m (CG9369, FBgn0002577)

Here, the many unannotated genes are particularly interesting to pursue.

Towards Aim 3 (b), measuring reversion according to a reporter of the cell growth pathway called S6, we identified 295 genes as novel regulators of S6 phosphorylation. Of these genes, 198 show low-pS6 RNAi phenotypes, 97 show high p-S6 phenotypes, and 197 genes (142 low-pS6, 55 high-pS6) encode proteins conserved in humans. Several of the conserved hits have wholly unknown functions, and we will look into their mechanisms of action. The genes we identified in this screen may represent drug targets for TSC.

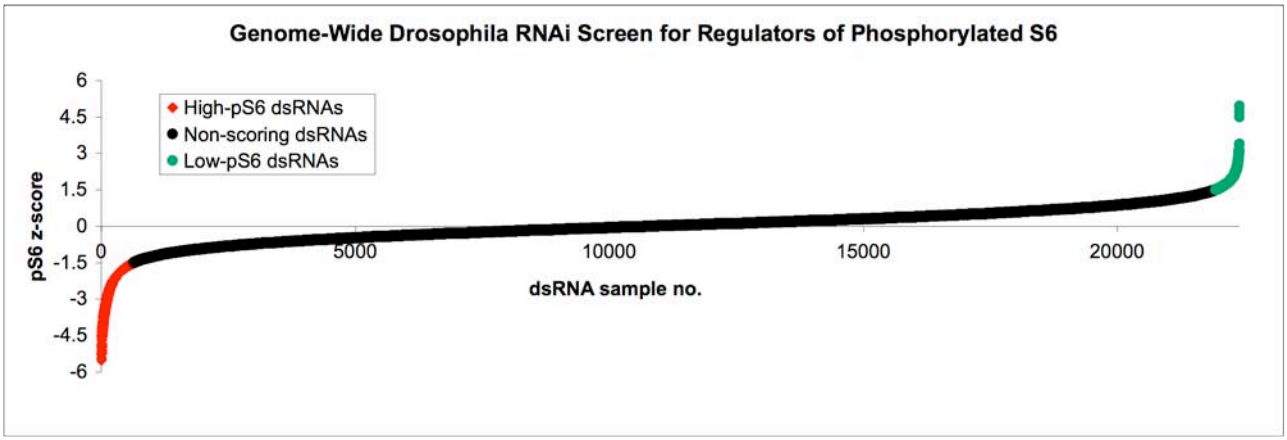

Figure 7: Overview of results from the p-S6 genome-wide screen. Samples are rank-ordered from high-p-S6 at the left to low-p-s6 at the right. Samples marked red and green meet the criteria as high and low hits, respectively.

Of the hits in this screen, the following have been validated for their effects on p-S6 levels by western blotting:

- Low hits: CG2028 (CKIalpha), CG4003 (pontin), CG9750 (reptin), CG2168 (RpS3a), CG3751 (RpS24), CG8857 (RpS11), CG4097 (Pros26, aka Proteasome B1 Subunit), CG17331 (aka Proteasome B2 Subunit), CG12323 (Prosbeta5, aka Proteasome B5 Subunit), CG4904 (Pros35, aka Proteasome A1 Subunit), CG5289 (Pros26.4, aka Proteasome C1 Subunit), CG11888 (Rpn2, aka Proteasome D1 Subunit)
- High hits: CG10944 (RpS6), CG6684 (RpS25)

We have begun low-throughput experiments to confirm these hits and identify their mechanisms of action. We have successfully confirmed the first 15 genes that we tested. From our first followup experiments, our most interesting discoveries so far include:

- Pontin and Reptin (Figure 8), which are well-conserved in mammals and are suspected to regulate cell growth through yet-unclear mechanisms. We confirmed the pS6 phenotype by Western blotting and determined also that Pontin and Reptin regulate pS6K in Drosophila cell lysates. We have some evidence that Pontin and Reptin directly bind to TOR in mammalian cells, and our lab will continue study of these proteins as potential TSC-independent regulators of TOR Complex 1.

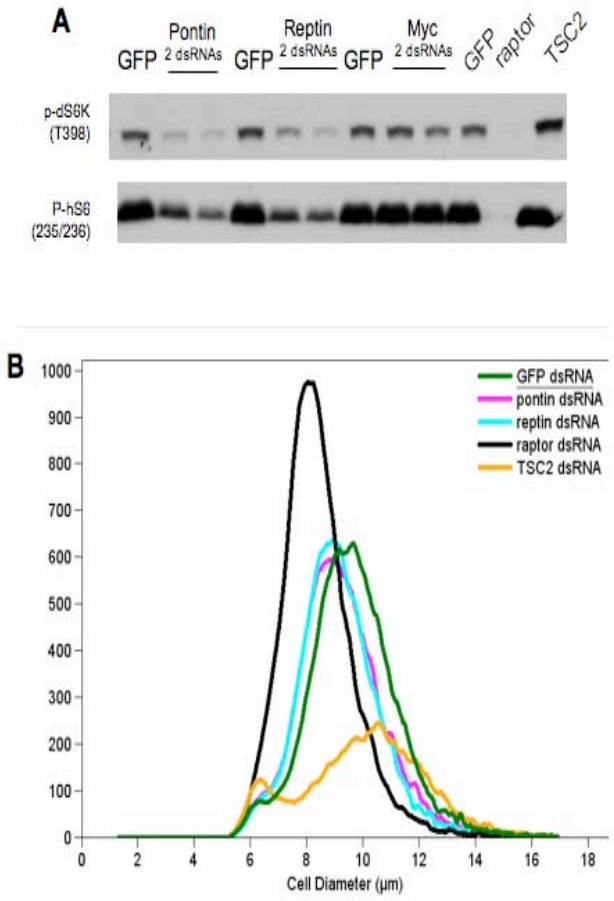

Figure 8: Validation of Pontin and Reptin as regulators of pS6. A) Pontin and Reptin were each knocked down with two non-overlapping dsRNAs. The oncogene Myc, which has been proposed to mediate Pontin and Reptin function (Bellosta et al, PNAS, 2005) was also tested as a control. Cell lysates were probed for pS6 and pS6K by Western blotting. B) S2 cells were transfected with the indicated dsRNAs and cell diameter was measured on a Coulter Counter.

- Casein Kinase I Alpha (Figure 9), which scored as a significant hit for low pS6 and small cell size. We confirmed the size phenotype with a Coulter counter, and we confirmed that CKIalpha RNAi affects not only the amount of phosphoS6 but also phosphoS6K in Western blots of Drosophila cell lysates. Furthermore we were able to show that CKIalpha RNAi's effects on Drosophila cells are independent of TSC loss.

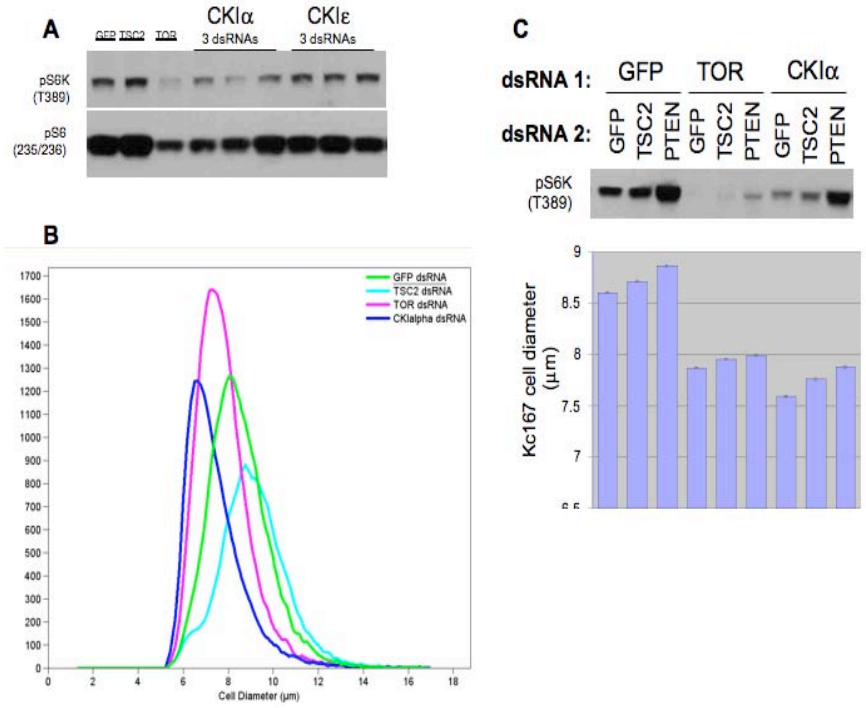

Figure 9: Validation of Casein Kinase I alpha as a regulator of pS6. A) CKIalpha was knocked down with three non-overlapping dsRNAs. The closely related kinase CKIepsilon was also tested as a control. S6 reporter S2R+ cell lysates were probed for pS6 and pS6K by Western blotting. B) Kc167 cells were transfected with the indicated dsRNAs and cell diameter was measured on a Coulter counter. C) Kc167 cells were transfected for 2 days with GFP/TOR/CKIalpha dsRNAs, and subsequently with dsRNAs targeting GFP/TOR/CKIalpha. Cell lysates were probed for pS6K by Western blotting (top panel), and mean cell diameter was measured on a Coulter counter (bottom panel). Error bars indicate 95% confidence levels.

- Many proteasomal subunits, including all the catalytic subunits and about two-thirds of the regulatory subunits (Figure 10). All these proteasomal hits scored low for pS6, which we confirmed by Western blot experiments that also demonstrated a low pS6K phenotype. We are performing follow-up experiments to determine the precise mechanism by which the proteasome regulates S6K phosphorylation.

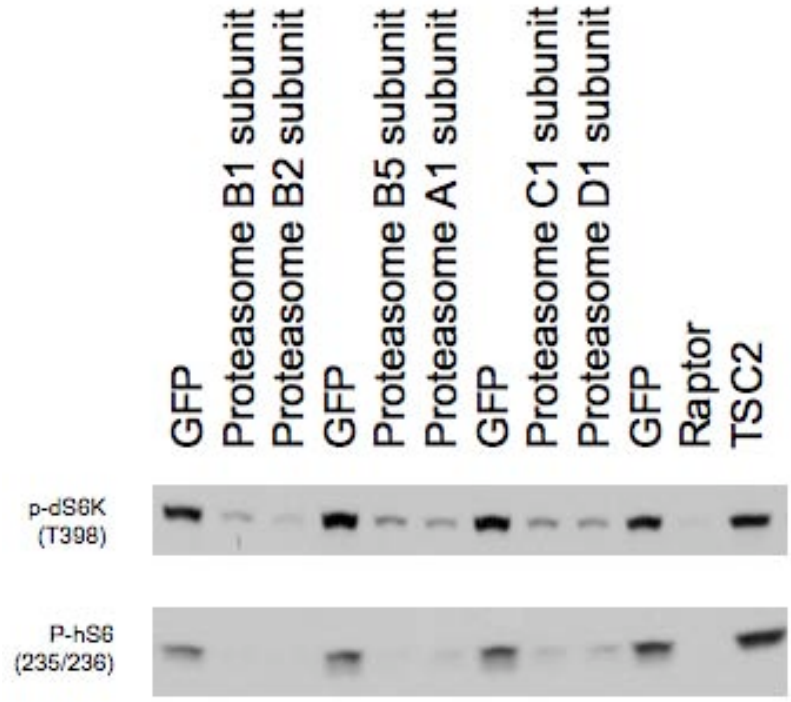

Figure 10: Validation of Proteasomal subunits as regulators of pS6. Various subunits of the proteasome were knocked down and cell lysates were probed for pS6 and pS6K by Western blotting.

Additional Results, beyond the Aims proposed:

As described in the introduction and the proposal, disruption of TSC1 and TSC2 have clinically and experimentally proven to yield similar phenotypes. By conducting genome-wide screens, we had the rare opportunity to test the similar function of these two genes by testing TSC1 and TSC2 knockdowns in the context of every *other* gene in the genome being knocked down by RNAi. The results, shown for cell size and for cell count (Figure 11) indicate that indeed, these two genes are behaving similarly across a wide number of cell states. This provides further evidence that the two genes entire function is within the context of a complex of the two proteins.

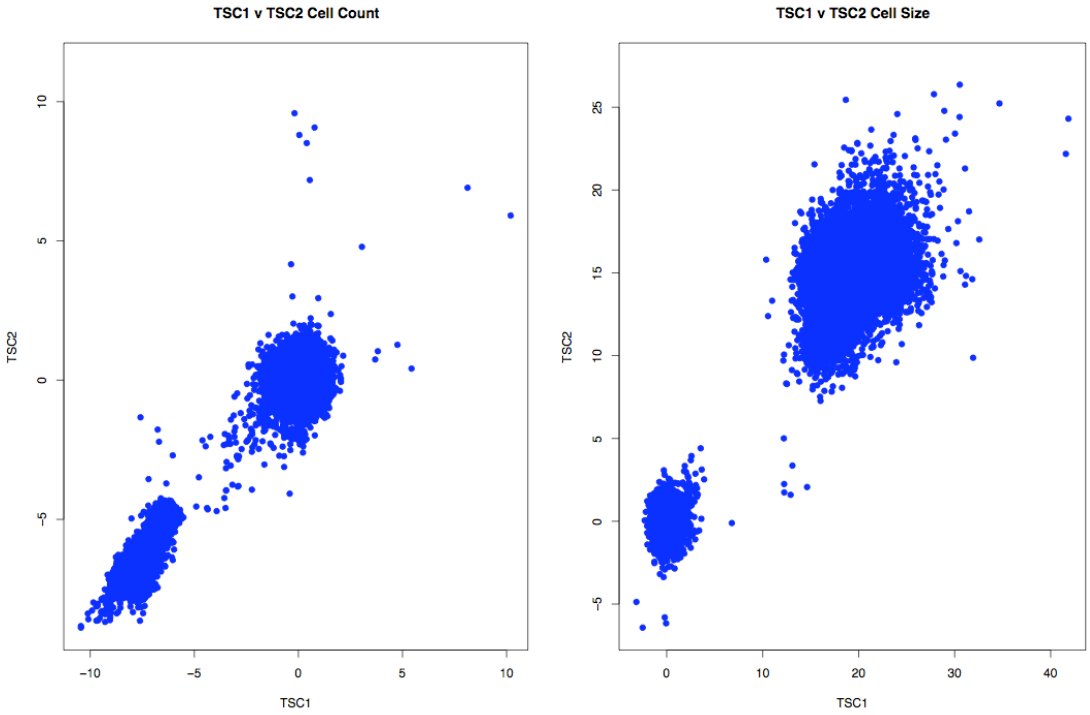

Figure 11: Each gene is plotted on the chart, with the median z-score for TSC1 on the horizontal axis and the median z-score for TSC2 on the vertical axis. Most genes are clustered along a diagonal line, indicating that TSC1 and TSC2 knockdowns respond similarly to various cellular states induced by knockdown of other genes.

Future plans inspired by this grant:

In addition to this low-throughput characterization of our most interesting hits, we intend to perform a followup array-based screen on all the hits from the wild-type p-S6 screen in order to determine their epistasis relationships with TSC1/2. As a minor point, we will also be interested to test our machine learning scoring system to score cells for apoptosis based on DNA and actin staining, to see whether this adds additional information to the screens completed. Lastly, before we determine precisely which genes to follow up on in future work, we would like to try several methods of statistical analysis, in particular combining the results from the separate TSC1 and TSC2 genome-wide screens, to obtain more statistical confidence.

Persons paid from this grant:

David M. Sabatini
Michael Lamprecht
Ola Friman

Key Research Accomplishments

- The dsRNA printing technology was developed, improved, and scaled up to high-throughput such that we were able to print ~20,000 dsRNAs, covering the vast majority of Drosophila genes, onto four glass microscope slides. In particular, technical difficulties related to the unwanted spreading of dsRNA in the spots and the imprecision of the DNA microarraying robot were resolved.
- We determined that software we wrote, CellProfiler, produces accurate quantitative cell measurements and is another key enabling technology now in place to analyze genome-wide screens.
    - Algorithms for the successful identification of Drosophila nuclei and cells in images were developed and validated.
    - We added the ability to measure a large number of sophisticated measurements to CellProfiler, including many measures of the size, shape, intensity, and texture of cells.
    - CellProfiler was validated for many key phenotypes. We tested its ability to measure cell number, cell size, cell cycle distribution (based on DNA content), and amount of fluorescence per cell using cells with known variations in these features.
    - The user interface of CellProfiler was improved dramatically to eliminate some of the tedium of high-throughput image analysis and to allow non-experts the ability to conduct image analysis experiments, including documentation and a manual.
    - CellProfiler was adapted to make use of a cluster of computers, so that it can analyze images at a pace faster than image acquisition.
    - CellProfiler was adapted to export measurements to Excel and also to a database, a necessary feature for large-scale genome-wide experiments.
    - CellProfiler was beta-tested by several academic and industry research groups.
    - CellProfiler was released for free to the public, allowing further development by the open-source community.
    - During the extension period, we wrote up and published validation of CellProfiler in a high-profile publication[6].
- We determined that software we wrote, CellVisualizer, allows visualization and extraction of the measurements made by CellProfiler so that genome-wide screens can be rapidly analyzed and conclusions drawn.
- We developed machine-learning methods and incorporated them into CellVisualizer, allowing complex phenotypes to be scored using morphology markers (DNA and actin stain).
- We performed multiple genome-wide screens to complete the aims of the proposal.
- We followed up a number of hits from the screen in detail, performing biochemical and cell-based assays to confirm the genes involved in TSC1- and TSC2-related pathways and to characterize their function.
- This work has yielded a number of potential new genes in the TSC1/TSC2 pathways that await careful characterization.

Reportable Outcomes

Completed Manuscripts/Abstracts/Publications:

(all available at: http://www.broad.mit.edu/~anne/publications.htm**)**

- Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology, 7:R100. Freely available at: http://genomebiology.com/2006/7/10/R100
- Lamprecht MR, Sabatini DM, Carpenter AE (2007) CellProfiler: free, versatile software for automated biological image analysis. BioTechniques.
- Jones TR, Carpenter AE, Golland P, Sabatini DM (2006) Methods for high-content, high-throughput image-based cell screening. Proceedings of the Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB). Metaxas DN, Whitaker RT, Rittcher J, Sebastian T (Eds). Copenhagen, Denmark, October 5, pp 65-72.

Manuscripts in progress:

- Most of the work described in this report is currently being confirmed and compiled into a manuscript for publication, led by Robert Lindquist in the Sabatini lab.

Open-source software released to the public, funded in part by this grant:

- CellProfiler cell image analysis software (www.cellprofiler.org)

Presentations discussing this work:

David M. Sabatini (principal investigator):

- (talk) Pharmacology & Cancer Biology Signal Transduction Colloquium Seminar Series, Duke University, Feb. 2006, Durham, NC
- (talk) Center for Cancer Research, Mass General Hospital, Feb. 2005, Boston, MA
- (talk) Neuroscience Monday Seminars, March 2005, Children's Hospital, Boston, MA
- (talk) TargetTalk Meeting, March 2005, San Diego, CA
- (talk) Department of Pharmacology, University of Virginia, April 2005, Charlottesville, VA
- (talk) Pfizer Research Technology Center, June 2005, Cambridge, MA
- (talk) American Diabetes Association, June 2005, San Diego, CA
- (talk) GlaxoSmithKline, July 2005, Philadelphia, PA
- (talk) Cancer Models & Mechanisms Gordon Research Conf., July 2005, Smithfield, RI
- (talk) Protein Kinases and Protein P'n FASEB Meeting, July 2005, Snowmass, CO
- (talk) Glucose Transporter Biology FASEB Meeting, August 2005, Snowmass, CO
- (talk) Ariad Pharmaceuticals, August 2005, Cambridge, MA

Anne E. Carpenter (postdoctoral fellow):

- (talk) Open Microscopy Environment User's Meeting, March 2007, Paris, France
- (talk) Genomics Institute of the Novartis Foundation, November 2006, San Diego, CA
- (talk) Burnham Institute, November 2006, San Diego, CA
- (talk) Broad Institute Retreat, November 2006, Boston, MA
- (talk) Life Sciences Research Foundation Annual Meeting, October 2006, Baltimore, MD
- (talk) Stowers Institute, September 2006, Kansas City, MO
- (talk) Society for Biomolecular Screening Annual Meeting, September 2006, Seattle, WA

- (talk) Institute for Systems Biology, September 2006, Seattle, WA
- (talk) St. Jude's Children's Research Hospital, September 2006, Memphis, TN
- (talk) Novartis Institute for Biomedical Research, August 2006, Cambridge, MA
- (talk) Purdue University, August 2006, W. Lafayette, IN
- (talk) Northwestern University, August 2006, Evanston, IL
- (talk) Merck/Computational Systems Biology Initiative meeting, June 2006, Boston, MA
- (talk) International Society for Analytical Cytology Congress, May 2006, Quebec City, Canada
- (talk) Academic Strategies for High-Throughput and High-Content Pathway Analysis, May 2006, Houston TX
- (talk) Harvard Mini Symposium on Image Analysis, April 2006, Boston, MA
- (talk) Stowers Institute, February 2006, Kansas City, MO
- (talk) High-Content Analysis CHI, January 2006, San Francisco, CA
- (talk) Harvard Department of Systems Biology, November 2005, Boston, MA
- (talk) Cytometry Development Workshop, Asilomar, October 2005, Pacific Grove, CA
- (talk) Merck Automated Biotechnology group, October 2005, North Wales, PA
- (talk) MipTec Enabling Technologies for Drug Discovery, May 2005, Basel, Switzerland
- (talk) Roche, May 2005, Nutley, NJ
- (poster) Life Sciences Research Foundation Annual Meeting, Oct. 2005, Wash., DC
- (poster) Discovery on Target, October 2005, Boston, MA
- (poster) Whitehead Institute Annual Retreat, September 2005, Waterville Valley, NH
- (poster) Society Biomolecular Screening Annual Mtg, Sept. 2005, Geneva, Switzerland

Colin Clarke (undergraduate student):

- (talk) American Society for Cell Biologists Annual Mtg, Dec., 2005, San Francisco, CA

<u>Funding applied for based on this work:</u>

- Culpeper Biomedical Pilot Initiative grant
- L'Oreal for Women in Science fellowship
- Multiple contributions to larger grants (e.g., letters of support), not itemized here.
- NIH R01 application submitted to the National Cancer Institute: "Cell Growth Signaling in Cancer Development"

Conclusions

Through the funding of this project, using TSC1 and TSC2 as key signaling pathways of interest, Drosophila genome-wide RNA interference living cell microarrays have gone from a proof-of-principle concept to a robust technology. We have also developed and adapted a number of software tools to allow the analysis of these genome-wide image-based screens, a previously formidable challenge. We have completed the genome-wide experiments originally proposed, with modifications, and have thereby uncovered a number of genes involved in the TSC1/2 pathways, in particular those genes relating to the regulation of cell growth and proliferation. We have also completed substantial investigation of some of these genes so as to confirm them and characterize their importance in these signaling pathways. Much further work will be required to determine whether these genes will be useful drug targets for TSC and/or cancers.

References

[1] M. N. Hall, M. Raff, and G. Thomas, *Cell Growth: Control of Cell Size*. (Cold Spring Harbor Laboratory, 2004).

[2] Y. Li, M. N. Corradetti, K. Inoki et al., *Trends Biochem Sci* **29** (1), 32 (2004).

[3] D. B. Wheeler, S. N. Bailey, D. A. Guertin et al., *Nat Methods* **1** (2), 127 (2004).

[4] S. Armknecht, M. Boutros, A. Kiger et al., *Methods Enzymol* **392**, 55 (2005).

[5] T.R. Jones, A. E. Carpenter, and P. Golland, presented at the Proceedings of the ICCV Workshop on Computer Vision for Biomedical Image Applications (CVBIA), Beijing, China, 2005 (unpublished); T.R. Jones, A. E. Carpenter, D.M. Sabatini et al., presented at the Proceedings of the Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB), Copenhagen, Denmark, 2006 (unpublished).

[6] A. E. Carpenter, T. R. Jones, M. R. Lamprecht et al., *Genome Biol* **7** (10), R100 (2006).